



-62-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Chatterjee, Deb K.
- (ii) TITLE OF INVENTION: Cloned DNA Polymerases from *Thermotoga maritima* and Mutants Thereof
- (iii) NUMBER OF SEQUENCES: 47
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
  - (B) STREET: 1100 New York Ave., N.W., Suite 600
  - (C) CITY: Washington
  - (D) STATE: DC
  - (E) COUNTRY: USA
  - (F) ZIP: 20005-3934
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: US 09/229,173
  - (B) FILING DATE: 13-JAN-1999
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/706,702
  - (B) FILING DATE: 06-SEP-1996
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/689,807
  - (B) FILING DATE: 14-AUG-1996
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/537,400
  - (B) FILING DATE: 02-OCT-1995
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/576,759
  - (B) FILING DATE: 21-DEC-1995
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/537,397
  - (B) FILING DATE: 02-OCT-1995
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: US 08/525,057
  - (B) FILING DATE: 08-SEP-1995
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Millonig, Robert C.
  - (B) REGISTRATION NUMBER: 34,395
  - (C) REFERENCE/DOCKET NUMBER: 0942.2800008
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 202-371-2600

(B) TELEFAX: 202-371-2540

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Arg Xaa Xaa Xaa Lys Xaa Xaa Xaa Phe Xaa Xaa Xaa Tyr Xaa  
1 5 10

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2682 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..2679

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATG GCG AGA CTA TTT CTC TTT GAT GGC ACA GCC CTG GCC TAC AGG GCA	48
Met Ala Arg Leu Phe Leu Phe Asp Gly Thr Ala Leu Ala Tyr Arg Ala	
1 5 10 15	
TAT TAC GCC CTC GAC AGA TCC CTT TCC ACA TCC ACA GGA ATT CCA ACG	96
Tyr Tyr Ala Leu Asp Arg Ser Leu Ser Thr Ser Thr Gly Ile Pro Thr	
20 25 30	
AAC GCC GTC TAT GGC GTT GCC AGG ATG CTC GTT AAA TTC ATT AAG GAA	144
Asn Ala Val Tyr Gly Val Ala Arg Met Leu Val Lys Phe Ile Lys Glu	
35 40 45	
CAC ATT ATA CCC GAA AAG GAC TAC GCG GCT GTG GCC TTC GAC AAG AAG	192
His Ile Ile Pro Glu Lys Asp Tyr Ala Ala Val Ala Phe Asp Lys Lys	
50 55 60	
GCA GCG ACG TTC AGA CAC AAA CTG CTC GTA AGC GAC AAG GCG CAA AGG	240
Ala Ala Thr Phe Arg His Lys Leu Leu Val Ser Asp Lys Ala Gln Arg	
65 70 75 80	
CCA AAG ACT CCG GCT CTT CTA GTT CAG CAG CTA CCT TAC ATC AAG CGG	288
Pro Lys Thr Pro Ala Leu Leu Val Gln Gln Leu Pro Tyr Ile Lys Arg	
85 90 95	

CTG	ATA	GAA	GCT	CTT	GGT	TTC	AAA	GTG	CTG	GAG	CTG	GAG	GGA	TAC	GAA	336
Leu	Ile	Glu	Ala	Leu	Gly	Phe	Lys	Val	Leu	Glu	Leu	Glu	Gly	Tyr	Glu	
		100						105					110			
GCA	GAC	GAT	ATC	ATC	GCC	ACG	CTT	GCA	GTC	AGG	GCT	GCA	CGT	TTT	TTG	384
Ala	Asp	Asp	Ile	Ile	Ala	Thr	Leu	Ala	Val	Arg	Ala	Ala	Arg	Phe	Leu	
		115					120					125				
ATG	AGA	TTT	TCA	TTA	ATA	ACC	GGT	GAC	AAG	GAT	ATG	CTT	CAA	CTT	GTA	432
Met	Arg	Phe	Ser	Leu	Ile	Thr	Gly	Asp	Lys	Asp	Met	Leu	Gln	Leu	Val	
		130					135				140					
AAC	GAG	AAG	ATA	AAG	GTC	TGG	AGA	ATC	GTC	AAG	GGG	ATA	TCG	GAT	CTT	480
Asn	Glu	Lys	Ile	Lys	Val	Trp	Arg	Ile	Val	Lys	Gly	Ile	Ser	Asp	Leu	
		145			150					155					160	
GAG	CTT	TAC	GAT	TCG	AAA	AAG	GTG	AAA	GAA	AGA	TAC	GGT	GTG	GAA	CCA	528
Glu	Leu	Tyr	Asp	Ser	Lys	Lys	Val	Lys	Glu	Arg	Tyr	Gly	Val	Glu	Pro	
				165					170					175		
CAT	CAG	ATA	CCG	GAT	CTT	CTA	GCA	CTG	ACG	GGA	GAC	GAC	ATA	GAC	AAC	576
His	Gln	Ile	Pro	Asp	Leu	Leu	Ala	Leu	Thr	Gly	Asp	Asp	Ile	Asp	Asn	
			180					185					190			
ATT	CCC	GGT	GTA	ACG	GGA	ATA	GGT	GAA	AAG	ACC	GCT	GTA	CAG	CTT	CTC	624
Ile	Pro	Gly	Val	Thr	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Val	Gln	Leu	Leu	
		195					200					205				
GGC	AAG	TAT	AGA	AAT	CTT	GAA	TAC	ATT	CTG	GAG	CAT	GCC	CGT	GAA	CTC	672
Gly	Lys	Tyr	Arg	Asn	Leu	Glu	Tyr	Ile	Leu	Glu	His	Ala	Arg	Glu	Leu	
		210				215					220					
CCC	CAG	AGA	GTG	AGA	AAG	GCT	CTC	TTG	AGA	GAC	AGG	GAA	GTT	GCC	ATC	720
Pro	Gln	Arg	Val	Arg	Lys	Ala	Leu	Leu	Arg	Asp	Arg	Glu	Val	Ala	Ile	
		225			230					235				240		
CTC	AGT	AAA	AAA	CTT	GCA	ACT	CTG	GTG	ACG	AAC	GCA	CCT	GTT	GAA	GTG	768
Leu	Ser	Lys	Lys	Leu	Ala	Thr	Leu	Val	Thr	Asn	Ala	Pro	Val	Glu	Val	
				245					250					255		
GAC	TGG	GAA	GAG	ATG	AAA	TAC	AGA	GGA	TAC	GAC	AAG	AGA	AAA	CTA	CTT	816
Asp	Trp	Glu	Glu	Met	Lys	Tyr	Arg	Gly	Tyr	Asp	Lys	Arg	Lys	Leu	Leu	
			260					265					270			
CCG	ATA	TTG	AAA	GAA	CTG	GAG	TTT	GCT	TCC	ATC	ATG	AAG	GAA	CTT	CAA	864
Pro	Ile	Leu	Lys	Glu	Leu	Glu	Phe	Ala	Ser	Ile	Met	Lys	Glu	Leu	Gln	
		275					280					285				
CTG	TAC	GAA	GAA	GCA	GAA	CCC	ACC	GGA	TAC	GAA	ATC	GTG	AAG	GAT	CAT	912
Leu	Tyr	Glu	Glu	Ala	Glu	Pro	Thr	Gly	Tyr	Glu	Ile	Val	Lys	Asp	His	
		290				295					300					
AAG	ACC	TTC	GAA	GAT	CTC	ATC	GAA	AAG	CTG	AAG	GAG	GTT	CCA	TCT	TTT	960
Lys	Thr	Phe	Glu	Asp	Leu	Ile	Glu	Lys	Leu	Lys	Glu	Val	Pro	Ser	Phe	
		305			310					315					320	
GCC	CTG	GAC	CTT	GAA	ACG	TCC	TCC	TTG	GAC	CCG	TTC	AAC	TGT	GAG	ATA	1008
Ala	Leu	Asp	Leu	Glu	Thr	Ser	Ser	Leu	Asp	Pro	Phe	Asn	Cys	Glu	Ile	
				325					330					335		
GTC	GGC	ATC	TCC	GTG	TCG	TTC	AAA	CCG	AAA	ACA	GCT	TAT	TAC	ATT	CCA	1056
Val	Gly	Ile	Ser	Val	Ser	Phe	Lys	Pro	Lys	Thr	Ala	Tyr	Tyr	Ile	Pro	

B1

340	345	350	
CTT CAT CAC AGA AAC GCC CAC AAT CTT GAT GAA ACA CTG GTG CTG TCG Leu His His Arg Asn Ala His Asn Leu Asp Glu Thr Leu Val Leu Ser 355 360 365			1104
AAG TTG AAA GAG ATC CTC GAA GAC CCG TCT TCG AAG ATT GTG GGT CAG Lys Leu Lys Glu Ile Leu Glu Asp Pro Ser Ser Lys Ile Val Gly Gln 370 375 380			1152
AAC CTG AAG TAC GAC TAC AAG GTT CTT ATG GTA AAG GGT ATA TCG CCA Asn Leu Lys Tyr Asp Tyr Lys Val Leu Met Val Lys Gly Ile Ser Pro 385 390 395 400			1200
GTT TAT CCG CAT TTT GAC ACG ATG ATA GCT GCA TAT TTG CTG GAG CCA Val Tyr Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro 405 410 415			1248
AAC GAG AAA AAA TTC AAT CTC GAA GAT CTG TCT TTG AAA TTT CTC GGA Asn Glu Lys Lys Phe Asn Leu Glu Asp Leu Ser Leu Lys Phe Leu Gly 420 425 430			1296
TAC AAA ATG ACG TCT TAT CAG GAA CTG ATG TCG TTT TCC TCA CCA CTT Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Ser Pro Leu 435 440 445			1344
TTT GGT TTC AGC TTT GCG GAT GTT CCG GTA GAC AAG GCT GCC GAA TAC Phe Gly Phe Ser Phe Ala Asp Val Pro Val Asp Lys Ala Ala Glu Tyr 450 455 460			1392
TCC TGC GAG GAT GCA GAC ATC ACT TAT AGG CTC TAC AAG ATA CTC AGC Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Ile Leu Ser 465 470 475 480			1440
ATG AAG CTC CAT GAA GCG GAA CTT GAG AAC GTC TTC TAC AGG ATA GAG Met Lys Leu His Glu Ala Glu Leu Glu Asn Val Phe Tyr Arg Ile Glu 485 490 495			1488
ATG CCG TTG GTG AAC GTC TTG GCA CGA ATG GAA TTC AAC TGG GTG TAT Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Phe Asn Trp Val Tyr 500 505 510			1536
GTT GAC ACA GAA TTC CTG AAA AAG CTC TCG GAG GAG TAC GGC AAA AAG Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys 515 520 525			1584
CTC GAG GAA CTG GCC GAA AAA ATC TAC CAG ATA GCA GGT GAG CCC TTC Leu Glu Glu Leu Ala Glu Lys Ile Tyr Gln Ile Ala Gly Glu Pro Phe 530 535 540			1632
AAC ATC AAT TCT CCA AAA CAG GTT TCA AAC ATC CTT TTT GAG AAG CTG Asn Ile Asn Ser Pro Lys Gln Val Ser Asn Ile Leu Phe Glu Lys Leu 545 550 555 560			1680
GGA ATA AAA CCC CGT GGA AAA ACG ACA AAA ACA GGA GAT TAC TCT ACC Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr 565 570 575			1728
AGG ATA GAG GTG TTG GAA GAG ATA GCG AAT GAG CAC GAG ATA GTA CCC Arg Ile Glu Val Leu Glu Glu Ile Ala Asn Glu His Glu Ile Val Pro 580 585 590			1776

B1

CTC	ATT	CTC	GAG	TTC	AGA	AAG	ATC	CTG	AAA	CTG	AAA	TCG	ACC	TAC	ATA	1824
Leu	Ile	Leu	Glu	Phe	Arg	Lys	Ile	Leu	Lys	Leu	Lys	Ser	Thr	Tyr	Ile	
		595					600					605				
GAC	ACC	CTT	CCG	AAA	CTT	GTG	AAC	CCG	AAA	ACC	GGA	AGA	TTT	CAT	GCA	1872
Asp	Thr	Leu	Pro	Lys	Leu	Val	Asn	Pro	Lys	Thr	Gly	Arg	Phe	His	Ala	
		610				615					620					
TCT	TTC	CAC	CAG	ACG	GGT	ACC	GCC	ACT	GGC	AGG	TTG	AGT	AGC	AGT	GAT	1920
Ser	Phe	His	Gln	Thr	Gly	Thr	Ala	Thr	Gly	Arg	Leu	Ser	Ser	Ser	Asp	
					630					635					640	
CCA	AAT	CTT	CAG	AAT	CTT	CCG	ACA	AAG	AGC	GAA	GAG	GGA	AAA	GAA	ATT	1968
Pro	Asn	Leu	Gln	Asn	Leu	Pro	Thr	Lys	Ser	Glu	Glu	Gly	Lys	Glu	Ile	
				645						650				655		
AGA	AAA	GCG	ATT	GTG	CCC	CAG	GAT	CCA	GAC	TGG	TGG	ATC	GTC	AGT	GCG	2016
Arg	Lys	Ala	Ile	Val	Pro	Gln	Asp	Pro	Asp	Trp	Trp	Ile	Val	Ser	Ala	
			660					665					670			
GAT	TAT	TCC	CAA	ATA	GAA	CTC	AGA	ATC	CTC	GCT	CAT	CTC	AGT	GGT	GAT	2064
Asp	Tyr	Ser	Gln	Ile	Glu	Leu	Arg	Ile	Leu	Ala	His	Leu	Ser	Gly	Asp	
		675					680					685				
GAG	AAC	CTT	GTG	AAG	GCC	TTC	GAG	GAG	GGC	ATC	GAT	GTG	CAC	ACC	TTG	2112
Glu	Asn	Leu	Val	Lys	Ala	Phe	Glu	Glu	Gly	Ile	Asp	Val	His	Thr	Leu	
		690				695					700					
ACT	GCC	TCC	AGG	ATC	TAC	AAC	GTA	AAG	CCA	GAA	GAA	GTG	AAC	GAA	GAA	2160
Thr	Ala	Ser	Arg	Ile	Tyr	Asn	Val	Lys	Pro	Glu	Glu	Val	Asn	Glu	Glu	
					710					715					720	
ATG	CGA	CGG	GTT	GGA	AAG	ATG	GTG	AAC	TTC	TCT	ATA	ATA	TAC	GGT	GTC	2208
Met	Arg	Arg	Val	Gly	Lys	Met	Val	Asn	Phe	Ser	Ile	Ile	Tyr	Gly	Val	
				725					730					735		
ACA	CCG	TAC	GGT	CTT	TCT	GTG	AGA	CTT	GGA	ATA	CCG	GTT	AAA	GAA	GCA	2256
Thr	Pro	Tyr	Gly	Leu	Ser	Val	Arg	Leu	Gly	Ile	Pro	Val	Lys	Glu	Ala	
			740					745					750			
GAA	AAG	ATG	ATT	ATC	AGC	TAT	TTC	ACA	CTG	TAT	CCA	AAG	GTG	CGA	AGC	2304
Glu	Lys	Met	Ile	Ile	Ser	Tyr	Phe	Thr	Leu	Tyr	Pro	Lys	Val	Arg	Ser	
		755					760					765				
TAC	ATC	CAG	CAG	GTT	GTT	GCA	GAG	GCA	AAA	GAG	AAG	GGC	TAC	GTC	AGG	2352
Tyr	Ile	Gln	Gln	Val	Val	Ala	Glu	Ala	Lys	Glu	Lys	Gly	Tyr	Val	Arg	
		770				775					780					
ACT	CTC	TTT	GGA	AGA	AAA	AGA	GAT	ATT	CCC	CAG	CTC	ATG	GCA	AGG	GAC	2400
Thr	Leu	Phe	Gly	Arg	Lys	Arg	Asp	Ile	Pro	Gln	Leu	Met	Ala	Arg	Asp	
					790					795					800	
AAG	AAC	ACC	CAG	TCC	GAA	GGC	GAA	AGA	ATC	GCG	ATA	AAC	ACC	CCC	ATT	2448
Lys	Asn	Thr	Gln	Ser	Glu	Gly	Glu	Arg	Ile	Ala	Ile	Asn	Thr	Pro	Ile	
				805					810					815		
CAG	GGA	ACT	GCG	GCA	GAT	ATA	ATA	AAA	TTG	GCT	ATG	ATA	GAT	ATA	GAC	2496
Gln	Gly	Thr	Ala	Ala	Asp	Ile	Ile	Lys	Leu	Ala	Met	Ile	Asp	Ile	Asp	
			820					825					830			
GAG	GAG	CTG	AGA	AAA	AGA	AAC	ATG	AAA	TCC	AGA	ATG	ATC	ATT	CAG	GTT	2544
Glu	Glu	Leu	Arg	Lys	Arg	Asn	Met	Lys	Ser	Arg	Met	Ile	Ile	Gln	Val	

B1  
Cont

835		840		845	
CAT GAC GAA CTG GTC TTC GAG GTT CCC GAT GAG GAA AAA GAA GAA CTA					2592
His Asp Glu Leu Val Phe Glu Val Pro Asp Glu Glu Lys Glu Glu Leu					
850		855		860	
GTT GAT CTG GTG AAG AAC AAA ATG ACA AAT GTG GTG AAA CTC TCT GTG					2640
Val Asp Leu Val Lys Asn Lys Met Thr Asn Val Val Lys Leu Ser Val					
865		870		875	880
CCT CTT GAG GTT GAC ATA AGC ATC GGA AAA AGC TGG TCT TGA					2682
Pro Leu Glu Val Asp Ile Ser Ile Gly Lys Ser Trp Ser					
	885		890		

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 893 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Ala	Arg	Leu	Phe	Leu	Phe	Asp	Gly	Thr	Ala	Leu	Ala	Tyr	Arg	Ala
1				5					10					15	
Tyr	Tyr	Ala	Leu	Asp	Arg	Ser	Leu	Ser	Thr	Ser	Thr	Gly	Ile	Pro	Thr
			20					25					30		
Asn	Ala	Val	Tyr	Gly	Val	Ala	Arg	Met	Leu	Val	Lys	Phe	Ile	Lys	Glu
		35					40					45			
His	Ile	Ile	Pro	Glu	Lys	Asp	Tyr	Ala	Ala	Val	Ala	Phe	Asp	Lys	Lys
	50					55					60				
Ala	Ala	Thr	Phe	Arg	His	Lys	Leu	Leu	Val	Ser	Asp	Lys	Ala	Gln	Arg
65					70					75				80	
Pro	Lys	Thr	Pro	Ala	Leu	Leu	Val	Gln	Gln	Leu	Pro	Tyr	Ile	Lys	Arg
				85					90					95	
Leu	Ile	Glu	Ala	Leu	Gly	Phe	Lys	Val	Leu	Glu	Leu	Glu	Gly	Tyr	Glu
			100					105					110		
Ala	Asp	Asp	Ile	Ile	Ala	Thr	Leu	Ala	Val	Arg	Ala	Ala	Arg	Phe	Leu
	115						120				125				
Met	Arg	Phe	Ser	Leu	Ile	Thr	Gly	Asp	Lys	Asp	Met	Leu	Gln	Leu	Val
	130					135					140				
Asn	Glu	Lys	Ile	Lys	Val	Trp	Arg	Ile	Val	Lys	Gly	Ile	Ser	Asp	Leu
145					150					155					160
Glu	Leu	Tyr	Asp	Ser	Lys	Lys	Val	Lys	Glu	Arg	Tyr	Gly	Val	Glu	Pro
				165					170					175	
His	Gln	Ile	Pro	Asp	Leu	Leu	Ala	Leu	Thr	Gly	Asp	Asp	Ile	Asp	Asn
			180					185					190		

B1  
Conf

Ile Pro Gly Val Thr Gly Ile Gly Glu Lys Thr Ala Val Gln Leu Leu  
195 200 205

Gly Lys Tyr Arg Asn Leu Glu Tyr Ile Leu Glu His Ala Arg Glu Leu  
210 215 220

Pro Gln Arg Val Arg Lys Ala Leu Leu Arg Asp Arg Glu Val Ala Ile  
225 230 235 240

Leu Ser Lys Lys Leu Ala Thr Leu Val Thr Asn Ala Pro Val Glu Val  
245 250 255

Asp Trp Glu Glu Met Lys Tyr Arg Gly Tyr Asp Lys Arg Lys Leu Leu  
260 265 270

Pro Ile Leu Lys Glu Leu Glu Phe Ala Ser Ile Met Lys Glu Leu Gln  
275 280 285

Leu Tyr Glu Glu Ala Glu Pro Thr Gly Tyr Glu Ile Val Lys Asp His  
290 295 300

Lys Thr Phe Glu Asp Leu Ile Glu Lys Leu Lys Glu Val Pro Ser Phe  
305 310 315 320

Ala Leu Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asn Cys Glu Ile  
325 330 335

Val Gly Ile Ser Val Ser Phe Lys Pro Lys Thr Ala Tyr Tyr Ile Pro  
340 345 350

Leu His His Arg Asn Ala His Asn Leu Asp Glu Thr Leu Val Leu Ser  
355 360 365

Lys Leu Lys Glu Ile Leu Glu Asp Pro Ser Ser Lys Ile Val Gly Gln  
370 375 380

---

Asn Leu Lys Tyr Asp Tyr Lys Val Leu Met Val Lys Gly Ile Ser Pro  
385 390 395 400

Val Tyr Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro  
405 410 415

Asn Glu Lys Lys Phe Asn Leu Glu Asp Leu Ser Leu Lys Phe Leu Gly  
420 425 430

Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Ser Pro Leu  
435 440 445

Phe Gly Phe Ser Phe Ala Asp Val Pro Val Asp Lys Ala Ala Glu Tyr  
450 455 460

Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Ile Leu Ser  
465 470 475 480

Met Lys Leu His Glu Ala Glu Leu Glu Asn Val Phe Tyr Arg Ile Glu  
485 490 495

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Phe Asn Trp Val Tyr  
500 505 510

Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys  
515 520 525

B1  
cont

Leu Glu Glu Leu Ala Glu Lys Ile Tyr Gln Ile Ala Gly Glu Pro Phe  
530 535 540

Asn Ile Asn Ser Pro Lys Gln Val Ser Asn Ile Leu Phe Glu Lys Leu  
545 550 555 560

Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr  
565 570 575

Arg Ile Glu Val Leu Glu Glu Ile Ala Asn Glu His Glu Ile Val Pro  
580 585 590

Leu Ile Leu Glu Phe Arg Lys Ile Leu Lys Leu Lys Ser Thr Tyr Ile  
595 600 605

Asp Thr Leu Pro Lys Leu Val Asn Pro Lys Thr Gly Arg Phe His Ala  
610 615 620

Ser Phe His Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp  
625 630 635 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile  
645 650 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asp Trp Trp Ile Val Ser Ala  
660 665 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp  
675 680 685

Glu Asn Leu Val Lys Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu  
690 695 700

Thr Ala Ser Arg Ile Tyr Asn Val Lys Pro Glu Glu Val Asn Glu Glu  
705 710 715 720

Met Arg Arg Val Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val  
725 730 735

Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Ile Pro Val Lys Glu Ala  
740 745 750

Glu Lys Met Ile Ile Ser Tyr Phe Thr Leu Tyr Pro Lys Val Arg Ser  
755 760 765

Tyr Ile Gln Gln Val Val Ala Glu Ala Lys Glu Lys Gly Tyr Val Arg  
770 775 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp  
785 790 795 800

Lys Asn Thr Gln Ser Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile  
805 810 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Asp Ile Asp  
820 825 830

Glu Glu Leu Arg Lys Arg Asn Met Lys Ser Arg Met Ile Ile Gln Val  
835 840 845

His Asp Glu Leu Val Phe Glu Val Pro Asp Glu Glu Lys Glu Glu Leu  
850 855 860

B1  
cont.



Val Asp Leu Val Lys Asn Lys Met Thr Asn Val Val Lys Leu Ser Val  
865 870 875 880

Pro Leu Glu Val Asp Ile Ser Ile Gly Lys Ser Trp Ser  
885 890

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Pro Ser Phe Ala Leu Asp Leu Glu Thr Ser Ser  
1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Pro Val Phe Ala Phe Asp Thr Glu Thr Asp Ser  
1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Gly Pro Val Ala Phe Asp Ser Glu Thr Ser Ala  
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

B1

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Ile	Val	Ser	Asp	Ile	Glu	Ala	Asn	Ala
1				5					10

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GACGTTTCAA GCGCTAGGGC AAAAGA

26

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Arg	Arg	Val	Gly	Lys	Met	Val	Asn	Phe	Ser	Ile	Ile	Tyr	Gly
1				5						10			

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Arg Arg Ser Ala Lys Ala Ile Asn Phe Gly Leu Ile Tyr Gly  
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Arg Gln Ala Ala Lys Ala Ile Thr Phe Gly Ile Leu Tyr Gly  
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Arg Asp Asn Ala Lys Thr Phe Ile Tyr Gly Phe Leu Tyr Gly  
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly  
1 5 10

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GTATATTATA GAGTAGTTAA CCATCTTTCC A

31

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTAGGCCAGG GCTGTGCCGG CAAAGAGAAA TAGTC

35

(2) INFORMATION FOR SEQ ID NO:16:

- 
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 35 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAAGCATATC CTTGGCGCCG GTTATTATGA AAATC

35

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 694 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 2..691

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

G GAT CCA GAC TGG TGG ATC GTC AGT GCG GAT TAT TCC CAA ATA GAA Asp Pro Asp Trp Trp Ile Val Ser Ala Asp Tyr Ser Gln Ile Glu 895 900 905	46
CTC AGA ATC CTC GCT CAT CTC AGT GGT GAT GAG AAC CTT GTG AAG GCC Leu Arg Ile Leu Ala His Leu Ser Gly Asp Glu Asn Leu Val Lys Ala 910 915 920	94
TTC GAG GAG GGC ATC GAT GTG CAC ACC TTG ACT GCC TCC AGG ATC TAC Phe Glu Glu Gly Ile Asp Val His Thr Leu Thr Ala Ser Arg Ile Tyr 925 930 935 940	142
AAC GTA AAG CCA GAA GAA GTG AAC GAA GAA ATG CGA CGG GTT GGA AAG Asn Val Lys Pro Glu Glu Val Asn Glu Glu Met Arg Arg Val Gly Lys 945 950 955	190
ATG GTG AAC TTC TCT ATA ATA TAC GGT GTC ACA CCG TAC GGT CTT TCT Met Val Asn Phe Ser Ile Ile Tyr Gly Val Thr Pro Tyr Gly Leu Ser 960 965 970	238
GTG AGA CTT GGA ATA CCG GTT AAA GAA GCA GAA AAG ATG ATT ATC AGC Val Arg Leu Gly Ile Pro Val Lys Glu Ala Glu Lys Met Ile Ile Ser 975 980 985	286
TAT TTC ACA CTG TAT CCA AAG GTG CGA AGC TAC ATC CAG CAG GTT GTT Tyr Phe Thr Leu Tyr Pro Lys Val Arg Ser Tyr Ile Gln Gln Val Val 990 995 1000	334
GCA GAG GCA AAA GAG AAG GGC TAC GTC AGG ACT CTC TTT GGA AGA AAA Ala Glu Ala Lys Glu Lys Gly Tyr Val Arg Thr Leu Phe Gly Arg Lys 1005 1010 1015 1020	382
AGA GAT ATT CCC CAG CTC ATG GCA AGG GAC AAG AAC ACC CAG TCC GAA Arg Asp Ile Pro Gln Leu Met Ala Arg Asp Lys Asn Thr Gln Ser Glu 1025 1030 1035	430
GGC GAA AGA ATC GCG ATA AAC ACC CCC ATT CAG GGA ACT GCG GCA GAT Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp 1040 1045 1050	478
ATA ATA AAA TTG GCT ATG ATA GAT ATA GAC GAG GAG CTG AGA AAA AGA Ile Ile Lys Leu Ala Met Ile Asp Ile Asp Glu Glu Leu Arg Lys Arg 1055 1060 1065	526
AAC ATG AAA TCC AGA ATG ATC ATT CAG GTT CAT GAC GAA CTG GTC TTC Asn Met Lys Ser Arg Met Ile Ile Gln Val His Asp Glu Leu Val Phe 1070 1075 1080	574
GAG GTT CCC GAT GAG GAA AAA GAA GAA CTA GTT GAT CTG GTG AAG AAC Glu Val Pro Asp Glu Glu Lys Glu Glu Leu Val Asp Leu Val Lys Asn 1085 1090 1095 1100	622
AAA ATG ACA AAT GTG GTG AAA CTC TCT GTG CCT CTT GAG GTT GAC ATA Lys Met Thr Asn Val Val Lys Leu Ser Val Pro Leu Glu Val Asp Ile 1105 1110 1115	670
AGC ATC GGA AAA AGC TGG TCT TGA	694

B1  
Cont.

Ser Ile Gly Lys Ser Trp Ser  
1120

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 230 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Asp Pro Asp Trp Trp Ile Val Ser Ala Asp Tyr Ser Gln Ile Glu Leu  
1 5 10 15  
Arg Ile Leu Ala His Leu Ser Gly Asp Glu Asn Leu Val Lys Ala Phe  
20 25 30  
Glu Glu Gly Ile Asp Val His Thr Leu Thr Ala Ser Arg Ile Tyr Asn  
35 40 45  
Val Lys Pro Glu Glu Val Asn Glu Glu Met Arg Arg Val Gly Lys Met  
50 55 60  
Val Asn Phe Ser Ile Ile Tyr Gly Val Thr Pro Tyr Gly Leu Ser Val  
65 70 75 80  
Arg Leu Gly Ile Pro Val Lys Glu Ala Glu Lys Met Ile Ile Ser Tyr  
85 90 95  
Phe Thr Leu Tyr Pro Lys Val Arg Ser Tyr Ile Gln Gln Val Val Ala  
100 105 110  
Glu Ala Lys Glu Lys Gly Tyr Val Arg Thr Leu Phe Gly Arg Lys Arg  
115 120 125  
Asp Ile Pro Gln Leu Met Ala Arg Asp Lys Asn Thr Gln Ser Glu Gly  
130 135 140  
Glu Arg Ile Ala Ile Asn Thr Pro Ile Gln Gly Thr Ala Ala Asp Ile  
145 150 155 160  
Ile Lys Leu Ala Met Ile Asp Ile Asp Glu Glu Leu Arg Lys Arg Asn  
165 170 175  
Met Lys Ser Arg Met Ile Ile Gln Val His Asp Glu Leu Val Phe Glu  
180 185 190  
Val Pro Asp Glu Glu Lys Glu Glu Leu Val Asp Leu Val Lys Asn Lys  
195 200 205  
Met Thr Asn Val Val Lys Leu Ser Val Pro Leu Glu Val Asp Ile Ser  
210 215 220  
Ile Gly Lys Ser Trp Ser  
225 230

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Phe Leu Phe Asp Gly Thr  
 1 5

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 6 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Leu Leu Val Asp Gly His  
 1 5

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ser Leu Ile Thr Gly Asp Lys Asp Met Leu  
 1 5 10

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: not relevant  
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

*Bl  
cont*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Arg Ile Leu Thr Ala Asp Lys Asp Leu Tyr  
1 5 10

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 687 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCGTACCNGG GNTCNCNANA TCGACTGCAG CATGCAAGCT GGCTAATCAT GGTCATAGCT 60  
GTTTCCTGTG TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT 120  
AAAGTGTAAG GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC 180  
ACTGCCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA TCGGCCAACG 240  
CGCGGGGAGA GCGGGTTTGC GTATTGGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT 300  
GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT 360  
-----ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAAGGCC AGCAAAAAGGC 420  
CAGGAACCGT TAAAAAGGCC GCGTTGCTGG GCGTTTTTCC ATAGGCTCCG CCCCCCTTGA 480  
CGAGCATCAC AAAAATTCGA CGCTTCAAGT TCAGAGGTGG GCGAAACCCG ACAGGGGACTA 540  
TAAAGATTAC CAGGGCGTTT TCCCCCTGGG AAGCTNCCTT CGTGCGCTCT CCTGTTCCCG 600  
AACCTGGCCG GTTTAACCGG GATACCNGNT CGGCCTTTTN TCCCCTTNGG GGGAAANCCTT 660  
GGGNTTTTTN GNAAAANGCT AAGGGTT 687

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 701 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GCTCGTACCG GGGATCTNNN ANATCGACTG CAGCATGCAA GCTTGGCGTA ATCATGGTCA 60



TAGCTGTTTC CTGTGTGAAA TTGTTATCCG CTCACAATTC CACACAACAT ACGAGCCGGA	120
AGCATAAAGT GTAAAGCCTG GGGTGCCTAA TGAGTGAGCT AACTCACATT AATTGCGTTG	180
CGCTCACTGC CCGCTTTCCA GTCGGGAAAC CTGTCGTGCC AGCTGCATTA ATGAATCGGC	240
CAACGCGCGG GGAGAGGCGG TTGCGTATT GGGCGCTCTT CCGCTTCCTC GCTCACTGAC	300
TCGCTGCGCT CGGTCGTTTC GCTGCGGCGA GCGGTATCAG CTCACTCAAA GGCGGTAATA	360
CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA TGTGAGCAAA AGGCCAGCAA	420
AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG CTGGGCGTTT TTTCCATAGG CTCCGCCCCC	480
CTGANGAGCA TCANAAAAAT CGANGCTCAN GTCANAGGTG GCGAAACCCG ACAGGNCTAT	540
TAAAAGATNC CCAGGCGTTT TCCCCCTGG GAAGCTCCCT CGTGGGGCTC TCCTGGTTNC	600
GGNNCCCTGN CCGGNTTACC GGGGATAANC TTGTTCCGGN CTTTNTCCCC TTCNGGGAAA	660
ANGGTGGGGG GTTTTNTNNA AAAGGCTCAA AGGCTGGTAN G	701

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 717 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: both
  - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

---

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GNTNTAGNNN GGNCTAANNG GCGGGGAAAT CGAGCTCGGT ACCCGGGGAT CCTCTAGAGT	60
CGACCTGCAG GCATGCAAGC TTGGCGTAAT CATGGTCATA GCTGTTTCCT GTGTGAAATT	120
GTTATCCGCT CACAATTCCA CACAACATAC GAGCCGGAAG CATAAAGTGT AAAGCCTGGG	180
GTGCCTAATG AGTGAGCTAA CTCACATTAA TTGCGTTGCG CTCACTGCCC GCTTTCCAGT	240
CGGGAAACCT GTCGTGCCAG CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT	300
TGCGTATTGG GCGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCCGGC	360
TGCGGCGGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA GAAATCAGGG	420
GATAACGCAG GGAAAGAACA TGTGAGCAAA AGGCCAGCA AAAGGCCAGG AACCCGTAAA	480
AAGGCCGCGT TGCCTGGCGT TTTTCCATAG GCTCCGCCCC CCTTGACGAG CAATCACAAA	540
AATCGACGCT CAAAGTCAAG AGGTGGCGAA ACCCCGACAG GGAATTATAA AGATACCCAG	600
GCCGTTTCCC CCTGGAAGCT CCCCTCCGTG CGCTTCTCCT TGGTTCCCGA CCCTGCCGCT	660
TTACCNGGAT NCCTGTCCGC CCTTTTNTCC CTTTCNGGNA ACCGGGCGCT TTTTTTTT	717

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 713 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

NNNNNNNNNG GCTGANAGCG ATAAATCGAG CTCGGTACCC GGGGATCCTC TAGAGTCGAC	60
CTGCAGGCAT GCAAGCTTGG CGTAATCATG GTCATAGCTG TTCCTGTGT GAAATTGTTA	120
TCCGCTCACA ATTCCACACA ACATACGAGC CGGAAGCATA AAGTGTAAG CCTGGGGTGC	180
CTAATGAGTG AGCTAACTCA CATTAATTGC GTTGCCTCA CTGCCCCTT TCCAGTCGGG	240
AAACCTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACGC GCGGGGAGAG GCGGTTTGCG	300
TATTGGGCGC TCTTCGCTT CCTCGCTCAC TGACTCGCTG CGCTCGGTCG TTCGGCTGCG	360
GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA	420
CGCAGGAAAG AACATGTTGA GCAAAGGCC AGCAAAGGC CAGGAACCGT AAAAAGGCCG	480
CGTTTGCTGG CGTTTTTCCC ATAGGCTCCG CCCCCTTGA CGAACCATCA CAAAATCGA	540
CGCTCAATTC AGAAGTTGGC GAAAACCCGA CAGGACTAAT AAAGATACCC AGCGTTTCCC	600
CCCCTGGAAG CTCCCCTCCG TTGCGCCTCT CCCTGTTCCC GAACCTTGCC CGCTTACCGG	660
GAATACCTTG TCCNCCTTTT CTCCCCTTCC GGAANCGTT NGCGCCTTTC CCC	713

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Gly	Ser	Arg	Leu	Val	Asp	Arg	Gln	Cys	Gly	Leu	Phe	Pro	Asn	Arg	Thr
1				5				10						15	
Gln Asn Pro Arg Ser Ser Gln Trp															
20															

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 60 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Glu	Pro	Cys	Glu	Gly	Leu	Arg	Gly	Gly	His	Arg	Cys	Ala	His	Leu	Asp
1				5					10					15	
Cys	Leu	Gln	Asp	Leu	Gln	Arg	Lys	Ala	Arg	Arg	Ser	Glu	Arg	Arg	Asn
			20					25					30		
Ala	Thr	Gly	Trp	Lys	Asp	Gly	Glu	Leu	Leu	Tyr	Asn	Ile	Arg	Cys	His
		35				40						45			
Thr	Val	Arg	Ser	Phe	Cys	Glu	Thr	Trp	Asn	Thr	Gly				
	50					55					60				

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 99 amino acids  
    (B) TYPE: amino acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Arg	Ser	Arg	Lys	Asp	Asp	Tyr	Gln	Leu	Phe	His	Thr	Val	Ser	Lys	Gly
1			5						10					15	
Ala	Lys	Leu	His	Pro	Ala	Gly	Cys	Cys	Arg	Gly	Lys	Arg	Glu	Gly	Leu
			20					25					30		
Arg	Gln	Asp	Ser	Leu	Trp	Lys	Lys	Lys	Arg	Tyr	Ser	Pro	Ala	His	Gly
		35				40						45			
Lys	Gly	Gln	Glu	His	Pro	Val	Arg	Arg	Arg	Lys	Asn	Arg	Asn	Lys	His
	50					55					60				
Pro	His	Ser	Gly	Asn	Cys	Gly	Arg	Tyr	Asn	Lys	Ile	Gly	Tyr	Asp	Arg
65				70					75					80	
Tyr	Arg	Arg	Gly	Ala	Glu	Lys	Lys	Lys	His	Glu	Ile	Gln	Asn	Asp	His
			85						90					95	
Ser	Gly	Ser													

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Arg Thr Gly Leu Arg Gly Ser Arg  
1 5

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 6 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Gly Lys Arg Arg Thr Ser  
1 5

---

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Ser Gly Glu Glu Gln Asn Asp Lys Cys Gly Glu Thr Leu Cys Ala Ser  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

His Lys His Arg Lys Lys Leu Val Leu  
1 5

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Ile Gln Thr Gly Gly Ser Ser Val Arg Ile Ile Pro Lys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Asn Ser Glu Ser Ser Leu Ile Ser Val Val Met Arg Thr Leu  
1 5 10

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Arg Pro Ser Arg Arg Ala Ser Met Cys Thr Pro  
1 5 10

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Leu Pro Pro Gly Ser Thr Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Ser Gln Lys Lys  
1

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Thr Lys Lys Cys Asp Gly Leu Glu Arg Trp  
1 5 10

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Tyr Thr Val Ser His Arg Thr Val Phe Leu  
1 5 10

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Asp Leu Glu Tyr Arg Leu Lys Lys Gln Lys Arg  
1 5 10

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Leu Ser Ala Ile Ser His Cys Ile Gln Arg Cys Glu Ala Thr Ser Ser  
1 5 10 15

Arg Leu Leu Gln Arg Gln Lys Arg Arg Ala Thr Ser Gly Leu Ser Leu  
20 25 30

Glu Glu Lys Glu Ile Phe Pro Ser Ser Trp Gln Gly Thr Arg Thr Pro  
35 40 45

Ser Pro Lys Ala Lys Glu Ser Gln  
50 55

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Thr	Pro	Pro	Phe	Arg	Glu	Leu	Arg	Gln	Ile
1				5					10

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Glu	Lys	Glu	Thr
1			

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Ser	Phe	Arg	Phe	Met	Thr	Asn	Trp	Ser	Ser	Arg	Phe	Pro	Met	Arg	Lys
1				5					10					15	
Lys Lys Asn															

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide



*B1*  
*Abachide*  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Asn Ser Leu Cys Leu Leu Arg Leu Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Ala Ser Glu Leu Ala Gly Leu  
1 5

---